https://doi.org/10.1016/j.ymssp.2015.01.020

Locating Fatigue Damage Using Temporal Signal Features of Nonlinear Lamb Waves

Ming Hong[1,2#], Zhongqing Su[3,1*], Ye Lu[2], Hoon Sohn[4] and Xinlin Qing[5]

[1]Department of Mechanical Engineering
The Hong Kong Polytechnic University, Kowloon, Hong Kong

[2]Department of Civil Engineering
Monash University, Clayton, VIC 3800, Australia

[3]The Hong Kong Polytechnic University Shenzhen Research Institute
Shenzhen 518057, PR China

[4]Department of Civil and Environmental Engineering
Korea Advanced Institute of Science and Technology, Daejeon, 305-701, South Korea

[5]Department of Aviation Health and Safety Management
Beijing Aeronautical Science and Technology Research Institute of COMAC
Beijing, 102211, P.R. China

Submitted to *Mechanical Systems and Signal Processing*

(submitted on 25 August 2014; final version on 29 January 2015)

[#] PhD student

[*] Corresponding author. Tel.: +852 2766 7818; fax: +852 2365 4703. Email: zhongqing.su@polyu.edu.hk (Prof. Zhongqing Su, *Ph.D.*)

Abstract

The temporal signal features of linear guided waves, as typified by the time-of-flight (ToF), have been exploited intensively for identifying damage, with proven effectiveness in locating gross damage in particular. Upon re-visiting the conventional, ToF-based detection philosophy, the present study extends the use of temporal signal processing to the realm of nonlinear Lamb waves, so as to reap the high sensitivity of nonlinear Lamb waves to small-scale damage (*e.g.*, fatigue cracks), and the efficacy of temporal signal processing in locating damage. Nonlinear wave features (*i.e.*, higher-order harmonics) are extracted using networked, miniaturized piezoelectric wafers, and reverted to the time domain for damage localization. The proposed approach circumvents the deficiencies of using Lamb wave features for evaluating undersized damage, which are either undiscernible in time-series analysis or lacking in temporal information in spectral analysis. A probabilistic imaging algorithm is introduced to supplement the approach, facilitating a presentation of identification results in an intuitive manner. Through numerical simulation and then experimental validation, two damage indices (DIs) are comparatively constructed, based respectively on linear and nonlinear temporal features of Lamb waves, and used to locate fatigue damage near a rivet hole of an aluminum plate. Results corroborate the feasibility and effectiveness of using temporal signal features of nonlinear Lamb waves to locate small-scale fatigue damage, with enhanced accuracy compared with linear ToF-based detection. Taking a step further, a synthesized detection strategy is formulated by amalgamating the two DIs, targeting continuous and adaptive monitoring of damage from its onset to macroscopic formation.

***Keywords*:** temporal signal features; nonlinear Lamb waves; signal processing; fatigue damage; sparse sensor network; structural health monitoring

1. Introduction

Lamb waves, the elastic disturbance disseminating in a thin plate or shell-like structure, have been the subject of intense scrutiny over the years, based on which a diversity of nondestructive evaluation (NDE) and structural health monitoring (SHM) techniques have been deployed, in a cardinal effort to warrant the reliability, integrity and durability of aging engineering structures. Central to an increasing awareness of the use of Lamb waves is their appealing merits including the ability of promptly interrogating a large area with only a few transducers, the capacity to omni-directionally access hidden components, the high sensitivity to various types of damage, as well as the prospect for implementing *in-situ* SHM. The majority of present Lamb-wave-based NDE and SHM techniques exploit changes in the temporal signal features in the time domain, with respect to baseline signals, in the form of deviations in wave amplitude and/or phase. Of particular interests among the temporal signal features are time-of-flight (ToF) [1–3], wave reflections/transmissions [4, 5], energy dissipation [6], and mode conversions [7], to name a few.

In this backdrop, the theory and interpretation of temporal features of Lamb wave signals are prevalently based on the linear elasticity – extracting signal features at the frequency band at which the probing signals are generated. In that sense, the temporal features, for example the delay in ToF, show, to some extent, linear correlation with the alteration of material or structural parameters due to the damage. Thus, they are referred to as *linear temporal features* of Lamb waves in what follows, and the associated signal processing exercises as *temporal features processing*. In particular, ToF, the most straightforward yet informative linear temporal features, has proven effectiveness in locating gross damage (*viz.*, the damage with a characteristic dimension comparable to the wavelength of the probing waves) such as open cracks, through-holes, and voids [8, 9].

Yet, insofar as observed, the sensitivity of linear temporal features of Lamb waves is substantially restricted and wavelength-dependent. When used to characterize undersized damage, such as barely visible fatigue cracks or material degradation prior to the formation of discernable, macroscopic damage, these linear temporal features may become less sensitive. This is because inconspicuous damage (much smaller than the probing wavelength) would hardly alter linear temporal features and incur notable wave scattering phenomena. As a remedial measure, one can increase the excitation frequency of probing waves to achieve a reduced wavelength, but this is at the expense of introducing additional complexity to the signal appearance owing to the multimodal and dispersive properties of waves at higher frequencies. Therefore, when dealing with small-scale damage, linear temporal features of Lamb waves may compromise their effectiveness and accuracy.

As opposed to using linear temporal features, continued efforts have been casted to explore the nonlinear features of Lamb waves, with a hope to enhance the detectability of small-scale damage or even material degradation. When the probing Lamb waves traverse an elastic medium, the inherent nonlinearities of the medium and additional nonlinearities arising from possible damage can distort the probing waves. This results in a range of nonlinear attributes in the acquired Lamb wave signals, as evidenced at twice, thrice or higher-fold the probing frequency (a.k.a. fundamental frequency) – termed higher-order harmonics [10–14]; or at half of the probing frequency – comparatively called sub-harmonics [15]; or at mixed frequencies when another excitation (rather than the fundamental frequency) is used to modulate the probing waves (*e.g.*, spectral sidebands in nonlinear wave modulation spectroscopy) [16–19], *etc*. These nonlinear attributes can be locally intensified when the probing Lamb waves pass through the damaged region where the damage-induced nonlinearities exist. For example, according to the “breathing crack model” under cyclic loads [20], when a crack closes, compressive and shear stresses of propagating waves are transmitted through the crack; when the crack opens during dilation, waves are partially decoupled. These will jointly lead to a local nonlinearity widely recognized as the *contact acoustic nonlinearity* (CAN). The higher-order harmonics generated therein, especially the second-order harmonic (as the third- and higher-order harmonics are usually too weak in magnitude to be perceptible in the signals), have gained prominence in characterizing undersized fatigue damage [11–14]. As fatigue damage introduces such local nonlinearities in the material when interacting with probing waves, the magnitude of the second-order harmonics in a captured Lamb wave signal can accordingly serve as an indicator to the presence of fatigue damage in the monitored structure. In addition, as the mechanism of this kind of detection is based on nonlinear features of Lamb waves in the frequency domain, its effectiveness is, in principle, less restricted by the probing wavelength than using linear techniques. Input waves in a moderate frequency range can entertain the demand of the evaluation of small-scale damage.

However, to put it into perspective, identification, extraction, and interpretation of nonlinear features of Lamb waves in these approaches are usually implemented via a spectral analysis in the frequency domain, at the expense of losing temporal signal features such as ToF. Consequently, this creates an obvious barrier to reaching quantitative damage localization. Among a limited number of exceptions, Kim *et al.* [21] explored the spatial variation of the normalized acoustic nonlinearity parameter of longitudinal waves associated with various propagation paths of different distances to a fatigue fracture site. Such a correlation was then extended to Lamb waves later [22], accounting for a variety of wave propagation lengths and angles of incidence. Nevertheless, because this nonlinearity parameter would decrease rapidly to an uninformative level as the sensing path moves away from the damage site, this approach entails a dense sensor network configuration with sensors deliberately and strategically positioned. Thus, it diverged from the paradigm of SHM which preferably uses sparse sensor networks with minimum intrusion to the host structure. On the other hand, in order to perceive weak nonlinear features of Lamb waves, handheld bulky wedge ultrasonic transducers are usually used, to be manipulated in a narrow frequency band with a concentrated intensity. All of these have posed another challenge towards practical realization of *in-situ* SHM. Last but not least, nonlinear features of Lamb waves can be much more prone to environmental and instrumentation noise than their linear counterparts [14], raising concerns on their applicability in real-world scenarios. In quest of the possibility of using a built-in sensor network to acquire nonlinear properties of Lamb waves for small-scale fatigue damage detection, the authors of this paper [12, 14, 23] employed miniaturized piezoelectric wafers to form an active sensor network in a sparse configuration, with which the above-addressed correlation between the nonlinearity parameter and the distance from a sensing path to a fatigue damage site was investigated. This series of studies has extended damage characterization using nonlinear properties of Lamb waves to *in-situ* SHM.

To combine the respective merits of *temporal features processing* and *nonlinear features* of Lamb waves while circumventing their recognized limitations, in this study the nonlinear features of Lamb waves (*i.e.*, second-order harmonics), acquired via a sparse sensor network and extracted with a time-frequency analysis, are interpreted through ToF-based temporal features processing, in an attempt to achieve quantitative localization of small-scale fatigue damage. Two probability-based damage indices (DIs) are comparatively constructed using the linear and nonlinear features of Lamb waves, respectively, and are used to characterize fatigue damage near the rivet hole of an aluminum plate, in both numerical simulation and experiment. It is then a corollary to amalgamate the two Dis, in order to further shape a damage detection and monitoring strategy with a capacity of accommodating damage in different scales (from its onset to macroscopic formation). Instead of chunky wedge transducers, miniaturized lead zirconate titanate (PZT) wafers are adopted to form the sparse sensor network, as a critical step toward the implementation of *in-situ* and cost-effective SHM [24].

2. Nonlinear features of Lamb waves

Although nonlinear Lamb waves have been addressed in a considerable amount of literature, it is incumbent on us to recapitulate their key aspects related to damage detection, prior to the development of a DI based on the nonlinear Lamb waves. An elastic medium possesses, by nature, certain types of sources producing nonlinearities that can distort elastic waves guided by the medium, including mainly the medium material itself, damage-driven plasticity, geometric effect, loading conditions, breathing crack behaviors, hysteresis, frictional and thermal effects of crack faces, and so on. To exploit the nonlinear features associated with second harmonics, the probing Lamb waves are often excited at a monochromatic frequency (denoted by f_E). Should the probing waves be modulated by the nonlinearities of the medium and/or from the damage, additional spectral components would presumably appear at twice the excitation frequency, *i.e.*, $2f_E$, in the spectrum, which are the corresponding second harmonics of the probing waves. Identification of damage, at either qualitative or quantitative level, can be fulfilled by extracting and calibrating the second harmonics-related nonlinear features of acquired Lamb wave signals.

In its intact condition, the material exhibits some sort of weak mesoscopic nonlinearity over the entire volume [25]. As fatigue damage emerges, microstructure defects such as the dislocations accumulate under repetitive loads, and then form persistent slip bands that may nucleate micro-cracks, which, at the grain boundaries, produce micro-cracks on the scale of millimeters. Finally, micro-cracks coalesce and grow into macroscopic cracks that propagate through the material. This process results in a variety of localized, yet more pronounced nonlinear behaviors in the vicinity of fatigue damage site, strengthening the nonlinear effects of the probing waves. Detailed studies on the types of damage-related nonlinearities and the mechanisms of higher-order harmonic generation can be found elsewhere [25–30]. Thus, it can be assumed the damage-induced nonlinearities, reflected in the captured wave signals, can serve as a localized indication of fatigue damage.

The nonlinearity parameter, denoted by β in what follows, is a frequently used measure to calibrate the nonlinearity of Lamb waves and other ultrasonic waves alike. This parameter can be theoretically explained using the nonlinear stress-strain relation of a medium containing nonlinearity sources. It links the magnitude of the probing fundamental wave mode (Θ_1) and that of its paired second harmonic mode (Θ_2), according to

$$\beta = 8\Theta_2 / (\Theta_1^2 k^2 x) \cdot \gamma , \tag{1}$$

where k and x are the wavenumber and propagation distance, respectively; γ is a scaling function of Lamb waves, regardless of the presence of damage in the medium [10]. Stronger nonlinear effects around fatigue damage site is manifested by an increase in β.

Note that, due to the multimodal and dispersive natures of Lamb wave as illustrated by the theoretical dispersion curves shown in Fig. 1 (using aluminum as an example), only particular wave modes excited at prudentially selected frequencies could be exploited, so as to ensure a prominent, cumulative second harmonic generation. A rich body of research has gone into this issue and provided selection criteria for wave mode and excitation frequency. At a rudimentary level, the probing fundamental mode at f_E and its paired second harmonic mode at $2f_E$ should share roughly the same phase and group velocities, respectively, a condition termed *synchronism*, with non-zero power flux from f_E to $2f_E$ [27]. In line with this, the Lamb wave mode pair (S_1, S_2), as marked in Fig. 1, can be a candidate, of which the first-order symmetric mode S_1 is excited at a frequency-thickness product of circa 3.59 MHz·mm and travels at a group velocity of about 4,375 m/s, and S_2, the second-order symmetric mode at 7.18 MHz·mm, is the corresponding second harmonics of S_1 with the same phase and group velocity. Another advantage of using the (S_1, S_2) pair relies on that S_1 and S_2 feature the highest velocities, at f_E and $2f_{E,}$, respectively, so that they can be easily identified in the time domain.

3. Temporal processing for nonlinear features of Lamb waves

3.1. Principle of ToF-based temporal features processing

ToF is defined as the duration for a specific wave packet to travel a certain distance, and is one of the most representative yet straightforward temporal features to be retrieved from a Lamb wave signal. The attributes of ToF reflect, to some extent, a linear correlation with damage parameters such as its location. In conjunction with the use of a networked sensor array, ToF-based temporal features processing can be applied on time-series signals to facilitate damage localization. To illustrate the principle of ToF-based temporal features processing, a sensing path, $T_i - T_j$ ($i, j = 1, 2, \ldots, N, i \neq j$), from part of a sensor network composed by N PZT transducers is shown schematically in Fig. 2. Using Cartesian coordinates, the actuator T_i, the sensor T_j, and the center of the damage are supposed to be situated at (x_i, y_i), (x_j, y_j), and (x_d, y_d), respectively. A simple triangulation algorithm can be used to describe their relationship, as

$$\frac{\sqrt{(x_d-x_i)^2+(y_d-y_i)^2}}{V_i}+\frac{\sqrt{(x_d-x_j)^2+(y_d-y_j)^2}}{V_{d-s}}=t_{i-j}\,, \qquad (2)$$

where V_i and $V_{d\text{-}s}$ are the group velocities of the incident probing wave mode and the damage-scattered wave mode, respectively, which may or may not be equal, contingent upon whether mode conversion occurs. $t_{i\text{–}j}$ is the ToF for the incident probing wave to travel from the actuator to the damage, then to the sensor after damage scattering. Therefore, Eq. (2) mathematically depicts an elliptical (provided $V_i = V_{d\text{-}s}$) or an ellipse-like (provided $V_i \neq V_{d\text{-}s}$) locus, indicating all the possible damage locations perceived by sensing path T_i – T_j. If multiple sensing paths are available, as in a sensor network, the damage can be presumably located at the intersections of most loci.

3.2. ToF-based temporal features processing – linear damage index

In this study, in order to interpret linear features acquired by individual sensing paths of a sensor network, a damage index is established across the entire inspection area, with which the damage, if any, is expected to be "visualized" in an intuitive image. Here, the DI is defined in terms of the probability of damage occurrence at a particular spatial node in the inspection area, using a probabilistic imaging algorithm (PIA). PIA distinguishes itself from traditional damage imaging techniques such as Lamb wave tomography by using a much sparse transducer network and a faster image reconstruction algorithm [4, 31, 32]. In PIA, the inspected region is first meshed virtually with P×Q nodes (assuming the region is rectangular, without the loss of generality), and projected to an image with each image pixel corresponding to a spatial node in the inspection region. The DI at mesh point (x_m, y_n) (m = 1, 2, …, P; n = 1, 2, …, Q) perceived by sensing path T_i – T_j, denoted by

$DI_{i,j}(x_m, y_n)$, is then defined as

$$DI_{i,j}(x_m, y_n) = 1 - \int_{-z_{mn}}^{z_{mn}} f(z)dz, \tag{3}$$

where

$$f(z_{mn}) = \frac{1}{\sigma_{mn}\sqrt{2\pi}} \exp[-\frac{z_{mn}^2}{2\sigma_{mn}^2}]. \tag{4}$$

More specifically, $f(z_{mn})$ is the normal distribution function that relates the probability density of damage occurrence at mesh point (x_m, y_n) to z_{mn}, the shortest distance from that node to the locus; σ_{mn} is the standard deviation of the relevant damage feature as a tolerance factor in the imaging process, which can be obtained from experiments and adjusted empirically. Thus, the DI defined by Eq. (3) quantifies the probability of damage occurrence at each node, with respect to the locus obtained from every sensing path in the sensor network. It also implies that the closer one node is to the locus of a particular sensing path, the higher the probability of damage is at that mesh node. This raw DI depicts a source image that highlights all the possible damage locations determined by that sensing path defining such a locus.

Taking a step further by integrating information from all sensing paths in the network, the ultimate diagnostic image can be produced using an image fusion scheme based on the arithmetic mean of raw DIs, as

$$DI_{Genre} = \frac{1}{N(N-1)} \sum_{i,j=1,i\neq j}^{N} DI_{i,j}. \tag{5}$$

Here, the subscript “Genre” can be substituted by “L” (*i.e.*, DI_L) in the case that it is derived from ToF – the *linear temporal features* of Lamb waves. In the ultimate image, pixels with remarkably high field values would pinpoint the damage location, even the shape or orientation of the damage zone(s), to provide quantitative depiction of damage.

Nevertheless, it is noteworthy that this localization process, based on a linear damage index using ToF-based temporal features processing, is most effective only when the damage is gross enough, compared to the probing wavelength, so that the gross damage can warrant notable wave scattering and prominent changes in linear temporal features to be distinguished intelligibly in the time domain.

3.3. Temporal features processing for nonlinear Lamb waves – nonlinear damage index

Waking up to the advantages of temporal features processing for damage localization as well as its limitation towards detection of undersized damage, the above linear DI is "transplanted" into the domain of nonlinear Lamb waves, to make use of the high sensitivity of nonlinear Lamb waves to small-scale damage. In essence, efforts are made to extract temporal information of indicative events during wave propagation, not only at the fundamental frequency to explore conventional ToF information, but also at the corresponding double frequency to exploit nonlinear harmonics of Lamb waves.

More specifically, a selective mode, for example S_1, is excited at a central frequency of f_E through a PZT wafer in the sensor network (to meet the criteria of *synchronism* and *non-zero power flux* for cumulative harmonic generation as mentioned earlier); the propagating waves are then captured by the rest wafers in the network, in the benchmark (intact) and current (damaged) statuses, respectively. Subsequently, a short-time Fourier transform (STFT) is performed on all the signals to develop time-frequency spectrograms, from each of which two energy profiles are retrieved at f_E and $2f_E$, denoted by $E_{f_E}(t)$ and $E_{2f_E}(t)$, respectively. On one hand, an indicative event at $2f_E$, for instance an abrupt energy hump detected in $E_{2f_E}(t)$ in the current status, may be found deviating from its benchmark counterpart. Such an energy packet, identified as the S_2 mode, is generated exclusively when the incident S_1 wave is nonlinearly distorted by the damage, which acts as a new source that “emits” second harmonics at $2f_E$. On the other hand, $E_{f_E}(t)$ at f_E in the current status is anticipated to have trivial deviations from its benchmark counterpart, because undersized damage would not incur prominent changes in temporal features of Lamb waves at f_E, until the damage size approaches the wavelength of the fundamental mode to induce significant damage scattering. In short, constructing $E_{2f_E}(t)$ at $2f_E$ essentially translates damage information exposed in the frequency domain back into the time domain where the location information can be retrieved (a detailed signal processing procedure is to be elaborated in Section 4.3).

Figure 3 illustrates this proposed detection rationale using the selected mode pair (S_1, S_2), which derives itself from the framework of the linear technique as exhibited in Fig. 2, yet with distinct mechanism and substantially improved effectiveness in detecting undersized damage. Now, Eq. (2) can be readily applied to the nonlinear circumstance, in which $V_i = V_{d-s} = V_{S_1} = V_{S_2}$ (according to the criteria of *synchronism*). The subsequent steps would be identical to the linear features processing elaborated in Section 3.2.

Ultimately, this leads to a damage index based on nonlinear attributes of Lamb waves, which is comparatively denoted by DI_{NL}, to be computed using Eq. (5) with the subscript “Genre” replaced by “NL” this time, meaning a DI that uses ToF of nonlinear features of Lamb waves.

4. Proof-of-concept simulation

To examine the benefit of the proposed approach, the constructed linear and nonlinear DIs (DI_L and DI_{NL}, respectively) are applied comparatively on the data from finite element (FE) simulation first, to evaluate a fatigue crack, 1 mm in its length, in an aluminum plate.

4.1. FE model

The considered aluminum plate (6061-T6), measuring 450×300×3.18 mm^3, is shown schematically in Fig. 4(a). For convenience of discussion, a coordinate system is defined with its origin at the center of the plate. A rivet round hole with a diameter of 10 mm centered at (25 mm, 45 mm) is produced to serve, in an engineering context, as a fatigue crack initiator. The model is created in ABAQUS®/CAE and subsequently analyzed in ABAQUS®/EXPLICIT using a dedicated modeling technique developed by the authors previously [22]. In this modeling technique, the material, geometric, and plasticity-driven nonlinearities are modeled by a modified nonlinear stress-strain relation through a subroutine VUMAT. In particular, an extra local nonlinearity parameter is added to the global nonlinearity parameter in the vicinity of the fatigue damage (by calling a second set of property values in the same subroutine [22]), in order to reflect the plasticity-driven nonlinearity near fatigue damage site.

Four circular PZT wafers (denoted by T_i , i = 1, 2, 3, 4), 8 mm in diameter for each, are collocated on one side of the plate as actuators and sensors, forming an active sensor network and rendering six sensing paths. The 1-mm long fatigue crack is modeled as a through-thickness seam, running down from the bottom of the hole in *Y* direction with an initial clearance of zero between the two crack faces, with the in-plane midpoint of the crack approximately at (25 mm, 39.5 mm). The crack is enabled with the “breathing” behavior when the probing waves traverse it. Furthermore, a contact-pair interaction with associated properties is defined on the crack interface to supplement the modeling of CAN mentioned in Section 2. Three-dimensional eight-node brick elements in ABAQUS® (C3D8R), each sized at 0.2 mm in the in-plane dimensions, are used to mesh the plate, as displayed in Fig. 4(b). The PZT wafers are modeled by a thin disk made up of four elements, which are tied to the plate surface [33]. 15.5-cycle Hann-windowed sinusoidal tone bursts at f_E = 1,130 kHz are selected in accordance with the mode selection criteria detailed in Section 2, to excite the probing waves by imposing uniform in-plane radial displacements on the actuator's periphery. According to Fig. 1 and given the plate thickness of 3.18 mm, this particular frequency excites S_1 as the probing fundamental mode (which then enables cumulative second harmonic generation of S_2 for developing DI_{NL}). Structural responses are acquired by the other three sensors in the form of in-plane elemental strains. The above excitation-acquisition procedure is repeated on each wafer in turn so that twelve signals are collected from these six paths.

4.2. Linear DI results

Although there are no clear restrictions on the selection of the excitation frequency to construct DI_L in Eq. (5), a frequency would be considered appropriate if it is able to achieve good signal recognizability, reduced wave dispersion, concentrated energy, a small number of co-existing wave modes, and most importantly, sufficient sensitivity to the damage to be detected. In this regard, the frequency of 1,130 kHz (to be selected for constructing DI_{NL} later, according to the criteria of synchronism) gives rise to sound signal recognizability with well separated wave packets along all sensing paths.

As representative results, Fig. 5(a) shows the time-domain signals acquired via $T_1 - T_4$

from the plate, in both benchmark and current (fatigue damaged) statuses. As can be seen, there are no apparent damage-scattered waves in the current signal, due to the minute scale of the crack compared to the probing wavelength. This finding confirms that even at such a high frequency (1,130 kHz), the probing wavelength of S_1 (~3.9 mm) is still not small enough to characterize the crack at such a scale (~1 mm). As a result, extraction of linear signal features (*i.e.*, ToF) of damage-scattered wave component becomes vain. For this reason, a relatively large value for σ_{mn} (from Eq. (4)) is obtained, representing the lack of confidence in the chosen damage feature. With calculated DI_L using Eq. (5), the ultimate diagnostic image is shown in Fig. 5(b), which renders false diagnosis on the damage's location.

4.3. Nonlinear DI results

In parallel with the linear DI analysis, the same time-domain signals are reinvestigated. STFT is performed on all the signals in both benchmark and current statuses, as exemplified by the one shown in Fig. 6(a) for T_1 – T_4 in the current status (whose benchmark counterpart looks very similar at this stage). The window size for STFT is selected at 512 so as to achieve a compromise between the temporal and spectral resolutions.

The energy profiles, $E_{f_E}(t)$ at f_E and $E_{2f_E}(t)$ at $2f_E$ for both statuses, are plotted as a function of time, shown in Figs. 6(b) and 6(c), respectively. The almost identical profiles of $E_{f_E}(t)$ in Fig. 6(b) (dotted *vs.* solid), irrespective of the fatigue crack's occurrence, confirms the previous finding that no significant damage scattering has occurred in the time domain, because the scale of the damage is much smaller than the probing wavelength.

In the meantime, second harmonic features can be observed even in the benchmark profile in Fig. 6(c) (prior to the occurrence of fatigue damage), because the weak mesoscopic nonlinearity of the intact medium, as well as the mathematical nonlinearities involved in the FE program, also contribute to the generation of harmonics. Notably, however, none of these nonlinearities originates from the damage. These second harmonics are found at any excitation frequency whatsoever, and thus they are deemed irrelevant to the cumulative generation mechanism (i.e., $S_1 \rightarrow S_2$) described earlier.

Nevertheless, an extra hump can be clearly observed in $E_{2f_E}(t)$ at $2f_E$ for the damaged case, as shown by the dotted line in Fig. 6(c), deviating obviously from the benchmark (solid line), which can only be attributed to the presence of damage showing nonlinear traits. This particular packet is thus considered the S_2 mode generated as the paired second harmonics of S_1 upon its interaction with the fatigue damage, which arrives after the first-arrival S_1 as a result of the "detour", as illustrated schematically in Fig. 3.

Note that the first S_1 mode at f_E, arriving at t_1 in Fig. 6(b), is coupled by other nonlinear modes at $2f_E$ due to non-damage-related nonlinearities (*e.g.*, from the material and the FE program as explained earlier), which are contained in the first couple of packets arriving at t_1' as shown in Fig. 6(c). Theoretically, at $2f_E$, the fastest mode (S_2) is supposed to have the same velocity as S_1; however, a slight difference is observed between t_1 and t_1', which can be interpreted by the time smearing effect of STFT, especially when the size of the chosen time window is relatively large and the frequency of interest is high. This smearing effect tends to stretch out the tails of the energy profile of a wave packet with respect to its peak, which remains much steadier in time. To take this into account and to achieve a more accurate extraction of temporal features, the arrival time of the peak of wave packet, rather than its initial arrival time t_1 or t_1', is used, so is the case for identifying ToF for the damage-induced S_2. Now, following Eqs. (2) through (4), DI_{NL} can be calculated for this path using Eq. (5), based on which a raw source image is generated, as displayed in Fig. 6(d). Finally, by fusing individual source images obtained from all available sensing paths in the network, the ultimate diagnostic image is produced as Fig. 6(e). The imaging factor σ_{mn} is further adjusted so that a compromise is achieved between detection accuracy and noise tolerance. Furthermore, a threshold κ can be applied to reinforce the identification result. Here, κ is a preset percentage of the maximum field value of the ultimate image such that any field value less than the threshold level is forced to approach zero. Fig. 6(f) shows the improved image with κ = 90%, where the location of the fatigue crack is precisely highlighted and distinguished from the rivet hole. It is also relevant to emphasize that the highlighted damage region is corresponding to the fatigue crack initiated from the rivet hole, rather than the rivet hole itself. This result verifies the feasibility and effectiveness of the proposed localization algorithm using ToF-based temporal features processing of nonlinear Lamb waves.

5. Experimental validation

Subsequent to FE simulation, the proposed methodology is examined experimentally by evaluating a hairline fatigue crack in an aluminum plate with the same configuration as that in simulation.

5.1. Specimen and measurement set-up

Four PZT wafers, 8 mm in diameter for each, are surface-mounted on an intact aluminum (6061-T6) plate as photographed in Fig. 7(a), which is consistent with the one considered in FE simulation as given in Fig. 4(a). The wafers are instrumented to a signal generation and acquisition system using shielded wires. The same 15.5-cycle tone burst excitation, as used in the simulation, is applied on the actuator as the probing signal with a Tektronix 3000C arbitrary function generator at 20 $V_{p\text{-}p}$. Response signals are acquired by the other three wafers with a Tektronix 4034B mixed signal oscilloscope at a sampling rate of 100 MS/s with 512-time averaging.

After the benchmark testing, the specimen undergoes a high-cycle fatigue test, subject to a sinusoidal tensile load varying from 2 to 20 kN at 10 Hz, on an Instron® 8802 fatigue platform as photographed in Fig. 7(b). To facilitate the initiation of a fatigue crack, a tiny, yet sharp notch is inscribed at the bottom edge of the rivet hole as a stress riser. After roughly 1.2-million cycles, a barely visible fatigue crack through the plate thickness and in parallel with the 300-mm side is produced, measuring about 3 mm in length, as shown in Fig. 7(c). Note that the generated fatigue crack is deliberately longer than the one considered in FE simulation, in order to compare the respective applicability of the linear and nonlinear DIs (to be detailed in Section 6). The same signal excitation and acquisition procedure is implemented on the fatigued specimen.

5.2. Linear vs. nonlinear DIs

The experimental signals, after low-pass filtering, are processed with STFT. The amplitude profiles (*i.e.*, $E_{f_E}(t)$ at f_E and $E_{2f_E}(t)$ at $2f_E$) of two groups of representative signals (benchmark *vs.* current, via path $T_1 - T_4$) are displayed in Figs. 8(a) and 8(b), respectively.

It can be found that the profiles, even at the fundamental probing frequency (Fig. 8(a)), exhibit some deviation after the first packet (such a deviation is not clearly observed in the simulation, Fig. 5(a)). That is because the crack in the experiment, 3-mm long, is close enough to the probing wavelength of 3.9 mm, and under such a circumstance considerable damage scattering has already occurred even in the time history. By finding the ToF of the damage-scattered waves (from the time-domain signal, or alternatively from $E_{f_E}(t)$), DI_L is calculated for each sensing path, and all indices from individual paths are then fused to build the ultimate diagnostic image using a κ of 90%, as illustrated in Fig. 8(c), where the crack location is accurately identified. Contrasting this experimental result to the image obtained from simulation using linear DI (Fig. 5(b), in which the damage could not possibly be identified even with a threshold applied), the increased size of the crack does contribute, despite the different investigation methods, at least qualitatively to the improvement of the final diagnosis using DI_L.

In parallel with the linear method, Fig. 8(b) is scrutinized again to build DI_{NL}. Note that the smearing effect at the double frequency becomes stronger than in the FE case, as the size of the time window for experimental signals is increased. Nevertheless, the enhanced second harmonics at $2f_E$ totally stands out as a separate wave packet at t_2, in contrast to the benchmark profile, which enables a very punctual ToF extraction from $E_{2f_E}(t)$ and demonstrates the higher sensitivity of second harmonics to fatigue damage. Similar to the imaging procedure using linear DI (DI_L), the ultimate diagnostic image using DI_{NL} with κ = 90% is presented in Fig. 8(d), showing a good agreement with the reality. In fact, as damage size increases, macroscopic damage scattering strengthens, and the punctuality of ToFs from linear Lamb waves also improves, leading to the reduced disparity of imaging quality between the linear and nonlinear techniques. Note the degree of this punctuality is reflected in imaging as the sharpness of damage loci.

6. Synthesized DI and discussion

In the above, ToF-based temporal features processing is applied in parallel on the linear and nonlinear signal features of acquired Lamb waves. It has been found that, for smaller damage, DI_{NL} outperforms its linear counterpart DI_L by virtue of its higher sensitivity. However, this sensitivity roots in the fact that the energy level at the double frequency is significantly lower than that at the fundamental frequency in the first place. If the signal noise, especially at higher bands, increases to a considerable magnitude, as they probably would in real-world scenarios, useful second harmonics features may be inundated, leaving the nonlinear DI less credible. In this regard, linear techniques are more noise-tolerable with better adaptivity. It is thus a corollary to develop a synthesized damage detection and SHM strategy by combining the linear and nonlinear DI, so as to reap their respective merits, making the methodology potentially more effective to damage of various dimensions in changing ambient environments.

In this study, a synthesized damage index, DI_S, is developed using a weighted average of DI_L and DI_{NL} defined by Eq. (5), as

$$DI_S = w_L \cdot DI_L + w_{NL} \cdot DI_{NL}, \tag{6}$$

where w_L and w_{NL} are the weights of DI_L and DI_{NL}, respectively, which sum to 1. For a better resolution, DI_S can then be normalized against itself. The weights are assigned based on the stage of damage and/or the noise level. For a relatively healthy plate, a large w_{NL} is preferred at the beginning of the monitoring process in order to pinpoint any potential small-scale damage. If the crack, under continuous monitoring, keeps growing, or if ambient noise becomes significant, w_L can be gradually increased to improve the noise-tolerance of the diagnosis while maintaining its sensitivity.

Figure 9 provides three examples of the diagnostic results in simulation when the crack is small (1-mm long) and in the absence of noise, in which w_{NZ} = 0, 0.4, and 0.5, respectively. As can be seen, when the crack is too small to be accurately depicted solely by DI_L, small adjustments of the weights can make immediate improvement to the ultimate image using the synthesized DI. When w_{NL} = 0.5 and beyond, the damage can be indicated precisely.

To simulate the impact of noise on DI_S, different levels of white noise (up to 50 MHz) are added to the raw signals (benchmark and current) obtained from the experimental study, where the 3-mm crack is large enough to be characterized by both DI_L and DI_{NL}. The noise level is approximately 1–2% of the maximum amplitude of each signal, which is substantially larger than the original noise presented, if any. In the example of path $T_1 – T_4$, it is found that after band-pass filtering and STFT, the new amplitude profiles at the fundamental frequency is hardly impacted, as shown in Fig. 10(a), which would lead to a very similar ToF of damage-scattered waves as from Fig. 8(a). However, the new amplitude profiles at the double frequency start to deviate from each other considerably earlier, because signal noise has been increased to a level that is large enough to contaminate any useful second harmonics, which may otherwise have indicated the reinforcement of nonlinearity. Note that this noise contamination may either randomly increase or decrease the level of second harmonics of the current signal relative to the benchmark. Consequently, the ToF determined from Fig. 10(b) tends to be smaller than the value retrieved from Fig. 8(b), due to the early involvement of increased noise (after the first-arriving S_1). In such a situation, a synthesized DI may be utilized to improve the effectiveness and flexibility of the monitoring strategy.

Figure 11 further illustrate the diagnostic results from the noise-contaminated experimental signals, when w_{NL} = 1, 0.75, and 0.4, respectively. As predicted, a sole nonlinear analysis renders false diagnosis, as second harmonics are quite noise prone. As the weight of DI_L increases, the diagnosis quality of the image improves rapidly; at w_{NL} = 0.4, the crack has been able to be identified relatively accurately. Indeed, it is envisaged from the analysis above that challenges could still remain in the detection of emerging small-scale damage in an environment that is noisy in the first place, or under conditions with such as varying temperature that may adversely impact the performance of the technique. However, with the proposed DI_S a tradeoff can be manipulated between noise tolerance and detectability of small-scale damage to maximize applicability.

7. Conclusions

Temporal signal features (*e.g.*, ToF) of linear and nonlinear Lamb waves are investigated for small-scale fatigue damage localization in an aluminum plate, using a sparse PZT sensor network that is conducive to the implementation of *in-situ* SHM. Two damage indices, based respectively on linear and nonlinear temporal features of Lamb waves, are established through a probability-based diagnostic imaging algorithm. Case studies are conducted in FE simulation and experiments on fatigued aluminum plates bearing cracks of two different lengths. While the linear technique, using ToF of damage-scattered waves, is more noise-proof and effective with gross damage, it generally fails to identify a fatigue crack whose size is much smaller than the probing wavelength. ToF-based temporal features processing of nonlinear Lamb waves can greatly facilitate the localization of small-scale damage quantitatively. A synthesized DI is therefore proposed to combine the superior sensitivity of nonlinear Lamb waves to fatigue damage with the high noise-tolerance of linear technique, through which the weights of the individual linear and nonlinear DIs are mutually complemented, to cope with increasing crack size or varying noise level, enabling the approach for continuous, adaptive SHM in practice, in which the potential influences of crack growth and ambient conditions are of increasing concern.

Acknowledgments

This project is supported by the Hong Kong Research Grants Council via General Research Funds (No. 523313 and No. 15214414). This project is also supported by National Natural Science Foundation of China (Grant No. 51375414). The work is partially supported by Mid-career Researcher Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0017456). Ming Hong is grateful for the Endeavour Australia Cheung Kong Research Fellowship, and to Monash University, Australia.